

STATISTICAL SUPPLEMENT  
TO  
SUMMARY DESCRIPTION  
OF  
GRADE NINE LITERATURE  
OBJECTIVES, TEST ITEMS  
AND BLUEPRINT

(FIRST EDITION)



Prepared At The Direction Of  
HIGH SCHOOL EXAMINATIONS BOARD  
DEPARTMENT OF EDUCATION

CURRICULUM

PR  
87  
A33  
1968

ALTA  
328.04  
1968  
Gr9

CURRGDHT

CURR

EDMONTON, ALBERTA

March, 1968



828.4  
1968  
Gr9

**STATISTICAL SUPPLEMENT  
TO  
SUMMARY DESCRIPTION  
OF  
GRADE NINE LITERATURE  
OBJECTIVES, TEST ITEMS  
AND BLUEPRINT**

**(FIRST EDITION)**



**Prepared At The Direction Of  
THE JUNIOR HIGH SCHOOL EXAMINATIONS BOARD  
DEPARTMENT OF EDUCATION**

**EDMONTON, ALBERTA**

**March, 1968**



## ACKNOWLEDGEMENT

This booklet was prepared by a special committee appointed by the Junior High School Examinations Board. The committee consisted of the following members:

J. A. Wood .....	Ottewell Junior High School, Edmonton,      Chairman
R. A. Allers .....	Department of Education, Edmonton
J. A. Hornsby .....	Gilbert Paterson Junior High School, Lethbridge
Lawrence Radcliffe .....	Ritchie Junior High School, Edmonton
Robert Westbury .....	Edmonton Separate School Board, Edmonton
Mrs. S. A. White .....	Ponoka Junior High School, Ponoka

## FOREWORD

In 1967, a booklet entitled Summary Description of Grade Nine Literature Objectives, Test Items and Blueprint was issued to Alberta teachers. Therein, cognitive and affective objectives in the field of literature were classified under the categories of Knowledge, Comprehension, Higher Mental Processes, and the Affective Domain. Using sample test items drawn from the 1967 Literature Departmental Examination, the present supplement illustrates the categories set forth in the above mentioned booklet and, in addition, demonstrates the value of item analysis to the teacher who seeks to develop his skill in constructing, evaluating, and refining items for valid and reliable tests.

Each selected item, printed as it appeared in the examination booklet, is followed by item analysis data. The item is then evaluated first in terms of the thought processes involved in arriving at the correct answer and then in terms of the item analysis data. Because the data are presented exactly as they appear on a computer print-out, charts on pages 2 through 7 of this pamphlet demonstrate how to identify, interpret, and use the statistics included in item analysis.

The present document represents only one type of test from among the many approaches to evaluation effectively used by teachers in the classroom. Within the area of objective testing, multiple-choice questioning is receiving increased emphasis. The teacher can develop this aspect of his testing program by using this study to increase the accuracy, discriminatory power and validity of each test item that he uses.

## PART I

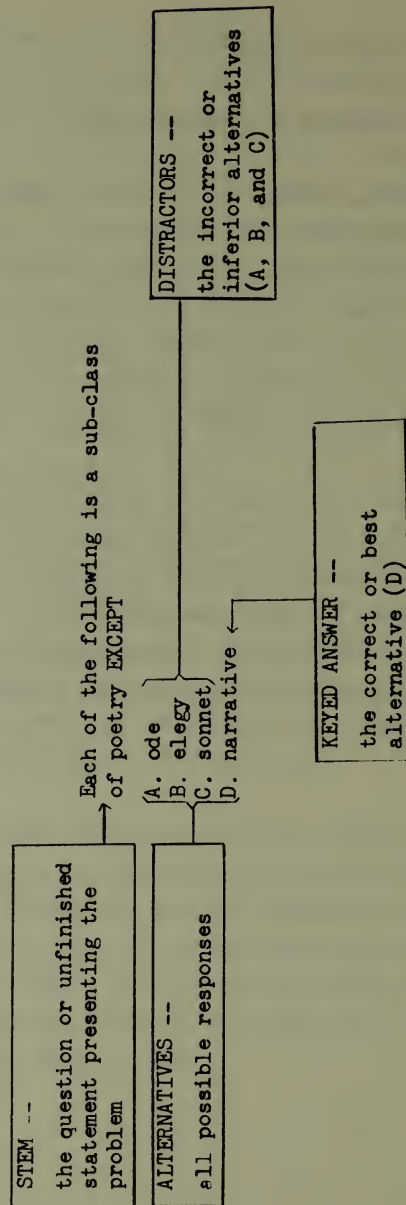
## ELEMENTS OF ITEM ANALYSIS

The item analysis data accompanying each item in PART II of this pamphlet could prove difficult for the teacher who is unfamiliar with the format in which the data is presented. Therefore, the following charts have been constructed to prepare the teacher for reading and using this data. Each chart introduces information that is complementary to that found in the preceding chart. Chart I labels and defines the parts of a multiple-choice item. Chart 2 and Chart 3 identify and define the elements of an item analysis. Chart 4 sets forth limits of acceptability for each element. Chart 5 illustrates an interpretation related to an acceptable item, while Chart 6 provides an interpretation related to an unacceptable item. It is recommended that the teacher carefully study the information and interpretations presented in these charts before proceeding to the detailed evaluations found in PART II.



# PARTS OF A MULTIPLE-CHOICE TEST ITEM

This chart identifies and explains terms as used in this booklet to describe the parts of a multiple-choice test item.





5. Total number of students choosing each alternative (Column under each total shows number of students contributing to that total.)

1, 2, 3, and 4 correspond to alternatives labelled A, B, C, and D in the test booklet. (Alternatives 1, 3, and 4 - in this item - serve as distractors.)

KEYED ANSWER to item

2. DIFFICULTY INDEX

1. NUMBER of students writing the test

Item number in test booklet

3. BISERIAL CORRELATION

4. ITEM RELIABILITY INDEX

TEST SCORE MEANS

8. TEST SCORE MEANS

6. UPPER TWO-FIFTHS

7. LOWER TWO-FIFTHS

5. Total number of students choosing each alternative (Column under each total shows number of students contributing to that total.)

1, 2, 3, and 4 correspond to alternatives labelled A, B, C, and D in the test booklet. (Alternatives 1, 3, and 4 - in this item - serve as distractors.)

KEYED ANSWER to item

2. DIFFICULTY INDEX

1. NUMBER of students writing the test

Item number in test booklet

3. BISERIAL CORRELATION

4. ITEM RELIABILITY INDEX

TEST SCORE MEANS

8. TEST SCORE MEANS

6. UPPER TWO-FIFTHS

7. LOWER TWO-FIFTHS

5. Total number of students choosing each alternative (Column under each total shows number of students contributing to that total.)

\*\*\* In a less sophisticated interpretation of item analysis, the figure under OMIT (number omitting item), the figures in the column under NR (no response), the figure under NF (paper not finished), the figures relating to the Middle Group (Upper 3), and the statistics for the Z-Score Means may be disregarded. On this chart, these elements are identified by the double asterisk (\*\*).

## DEFINITIONS

6. UPPER TWO-FIFTHS  
'UPPER 5' and 'UPPER 4'  
(60 to 100%-ile) include  
students receiving  
highest scores on total  
test. Sum of two num-  
bers representing UPPER  
TWO-FIFTHS in the column  
under each alternative  
is the figure used for  
comparisons explained  
in CHART 4.

UPPER 3 (Middle Group)  
40 to 60th %ile.  
Disregard

7. LOWER TWO-FIFTHS  
'UPPER 2' and 'UPPER 1'  
(0 to 40th %ile  
include students  
receiving lowest scores  
on total test. Sum of  
two numbers represent-  
ing LOWER TWO-FIFTHS in  
the column under each  
alternative is the  
figure used for com-  
parisons explained in  
CHART 4.

8. TEST SCORE MEANS  
Mean (average) score  
(on the total test)  
of students choosing  
this alternative. (In  
this instance, of 205  
students)

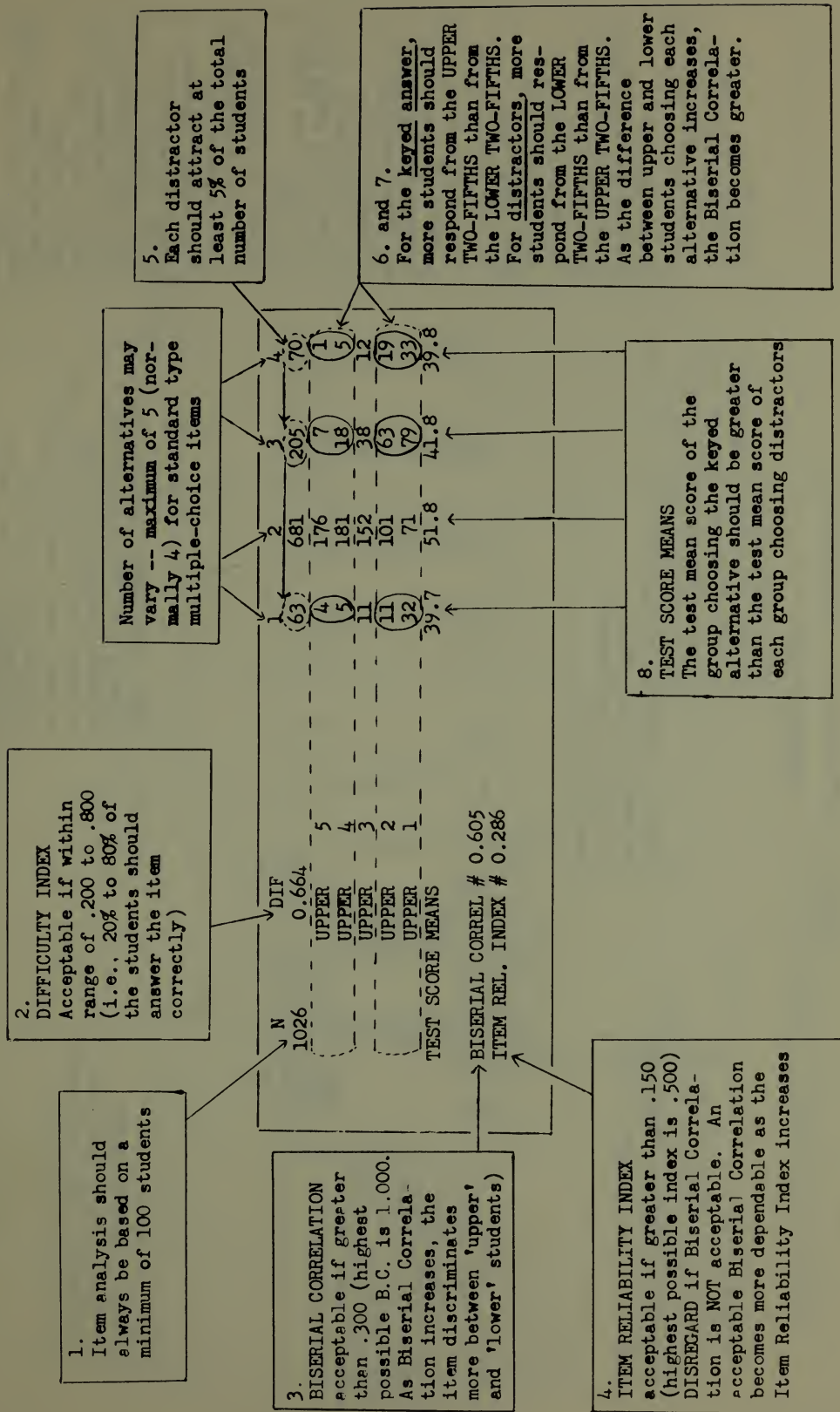
2. DIFFICULTY INDEX  
Proportion of students  
answering the item correctly

KEYED ANSWER - Correct answer  
to item (In this instance -  
alternative 2)

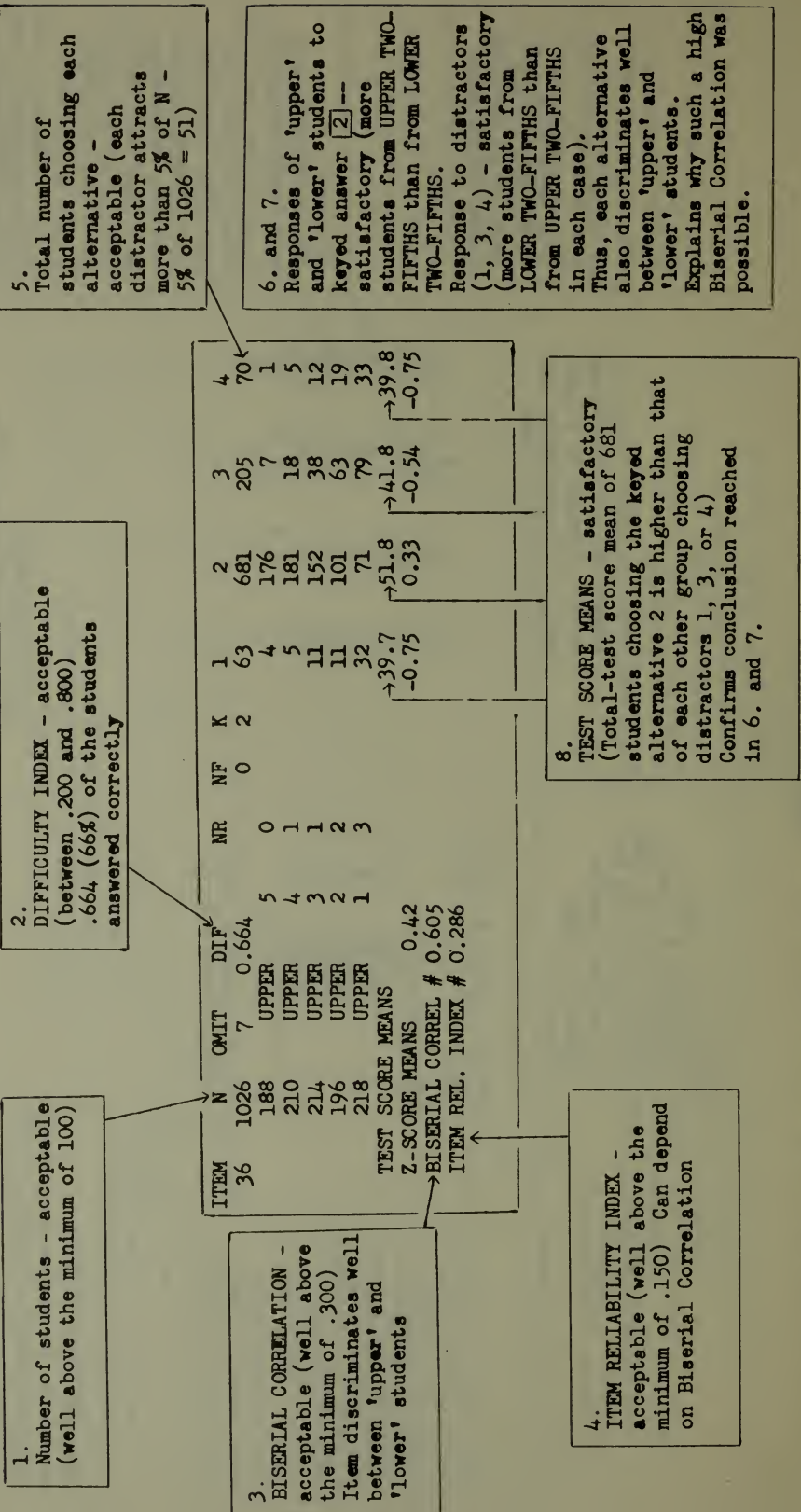
	K	1	2	3	4
DIF	2	1	2	3	4
	0.664			205	70
UPPER	5				1
UPPER	4				5
UPPER	3				12
UPPER	2				19
UPPER	1				33
MEANS				41.8	39.8

TEST SCORE MEANS

# ELEMENTS OF ITEM ANALYSIS ACCEPTABLE LIMITS



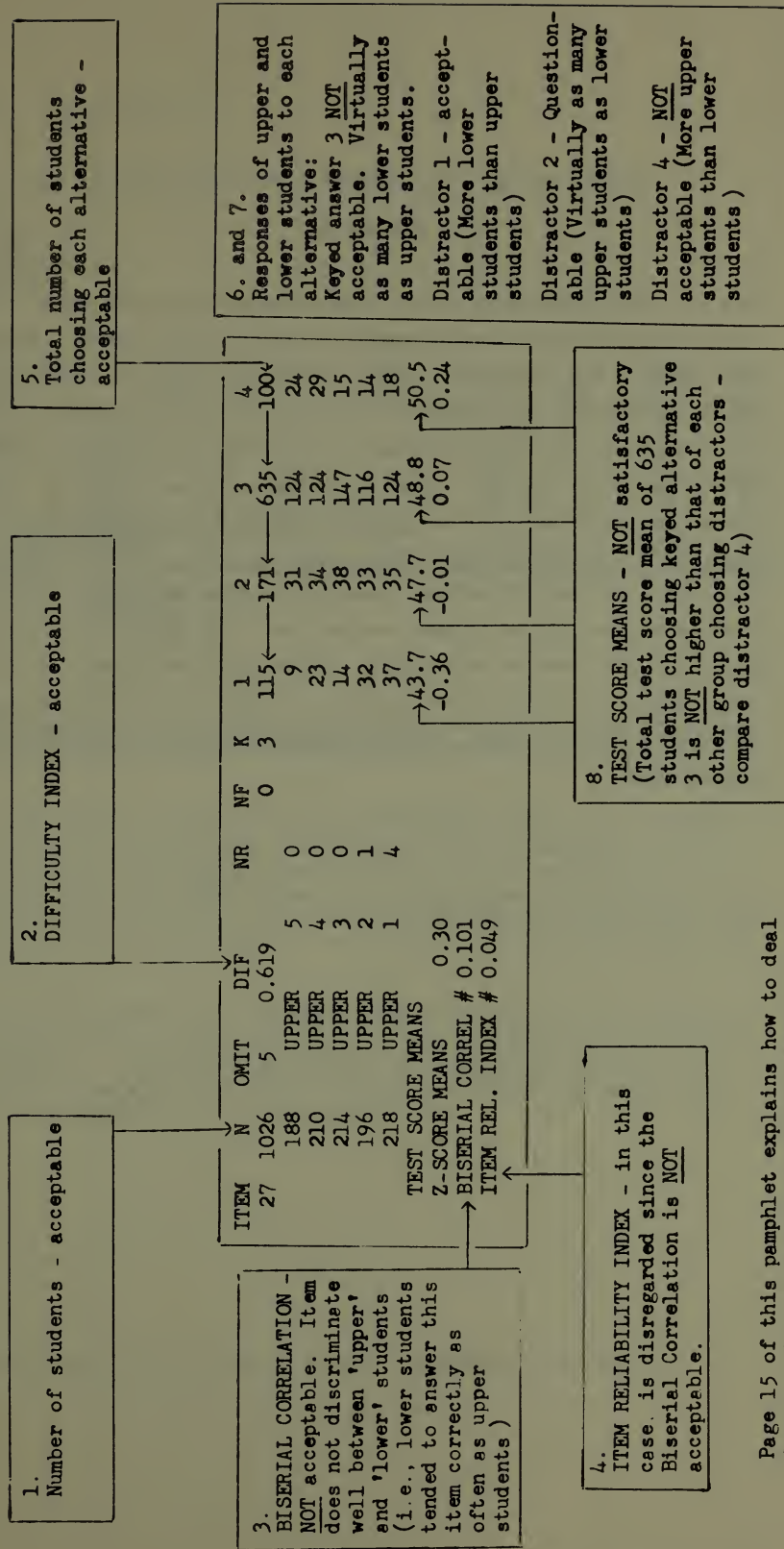
ELEMENTS OF ITEM ANALYSIS -- INTERPRETATION  
ACCEPTABLE ITEM





## ELEMENTS OF ITEM ANALYSIS -- INTERPRETATION

## UNACCEPTABLE ITEM



Page 15 of this pamphlet explains how to deal further with this initial interpretation.

PART II  
EVALUATIONS

The following evaluations have been compiled to reveal both strengths and weaknesses in sample multiple-choice items. Along with some of the successful items, unacceptable items were also selected from the 1967 Literature Departmental Examination in order to illustrate as many uses of the item analysis as possible. Effort has also been made to choose items representative of the categories set forth in the Summary Description of Grade Nine Literature Objectives, Test Items, and Blueprint, 1967. Complete statistics for each item have been included to facilitate a clearer comprehension of each evaluation. The general statistics - including the level of difficulty, Biserial Correlation, and the Item Reliability Index - are considered first. Then the detailed statistics are discussed. Where needed, suggestions for revisions are offered.

## EVALUATION 1 -- TEST ITEM 35

35. Each of the following is a sub-class of lyric poetry EXCEPT

- A. ode
- B. elegy
- C. narrative
- D. sonnet

ITEM	N	OMIT	DIF	NR	NF	K	1	2	③
35	1026	7	0.469				118	221	481
	188	UPPER	5	1	0	3	7	12	133
	210	UPPER	4	0			19	45	111
	214	UPPER	3	2			31	42	99
	196	UPPER	2	2			24	51	58
	218	UPPER	1	2			37	71	58
TEST SCORE MEANS							43.8	44.1	51.6
Z-SCORE MEANS							-0.36	-0.36	0.33
BISERIAL CORREL #									0.375
ITEM REL. INDEX #									0.187

This item was placed in the Knowledge category because the item requires the student to remember specific information about the classes of lyric poetry.

The level of difficulty indicates that approximately 47% of the students responded correctly to this item. This means that the question is of average difficulty. The Biserial Correlation of 0.375 reveals that the item discriminated well between those who scored higher on the total test and those who did not. The Item Reliability Index of 0.187 indicates that one can depend on this Biserial Correlation with a fairly high degree of confidence.

All distractors (incorrect alternatives) functioned very well. Number 2 was the strongest because it attracted 122 lower students and only 57 of the upper students. The weakest distractor is alternative 4 which was too attractive to the upper students. Revision of this item would probably require attention to this weak distractor.

\* In the statistics for each item to follow the Keyed Answer will be circled.



## EVALUATION 2 -- TEST ITEM 13

13. The selection which emphasizes the importance of the Bible in the upbringing of children is
- A. God is at the Anvil
  - B. My Mother's Voice
  - C. The Cotter's Saturday Night
  - D. The Happy Journey to Trenton and Camden

ITEM	N	OMIT	DIF	NR	NF	K	1	2	③	4
13	1026	171	0.438		0	3	115	110	449	181
	188	UPPER	5	43			2	4	115	24
	210	UPPER	4	31			14	11	119	35
	214	UPPER	3	40			24	14	97	39
	196	UPPER	2	29			27	42	60	38
	218	UPPER	1	28			48	39	58	45
TEST SCORE MEANS							40.9	41.5	51.4	46.7
Z-SCORE MEANS							-0.65	-0.54	0.24	-0.10
BISERIAL CORREL #							0.335			
ITEM REL. INDEX #							0.166			

This item was designed to test knowledge of the philosophy underlying a specific work. Since the student was required only to remember the theme of the correct alternative in order to answer the question correctly, this item was assigned to the Knowledge category.

The difficulty level, of approximately 44% indicates that the item was somewhat difficult. The Biserial Correlation and the Item Reliability Index show that the item discriminated quite acceptably between 'upper' and 'lower' students.

The detailed statistics indicate that each of the distractors was a plausible answer attracting more of the poorer students than the better students.

This item thus proved effective in its present form.

## EVALUATION 3 -- TEST ITEM 36

36. We all are blind until we see  
       That in the human plan  
 Nothing is worth the making if  
       It does not make the man.
- Why build these cities glorious  
       If man unbuilt goes?  
 In vain we build the world, unless  
       The builder also grows.

What is the poet's conclusion?

- A. We are all backward until we know both city and country life.  
 B. Character building should not be neglected.  
 C. Man grows when cities are built and grow large.  
 D. In the human plan, the urban population are supposed to grow.

ITEM	N	OMIT	DIF	NR	NF	K	1	(2)	3	4
36	1026	7	0.664		0	2	63	681	205	70
	188	UPPER	5	0			4	176	7	1
	210	UPPER	4	1			5	181	18	5
	214	UPPER	3	1			11	152	38	12
	196	UPPER	2	2			11	101	63	19
	218	UPPER	1	3			32	71	79	33
TEST SCORE MEANS							39.7	51.8	41.8	39.8
Z-SCORE MEANS							-0.75	0.33	-0.54	-0.75
BISERIAL CORREL # 0.605										
ITEM REL. INDEX # 0.286										

This item was designed to test the student's ability to interpret or to grasp the thought of a work as a whole. It has been placed in the Comprehension category.

The item was easy with approximately 66% of the students responding correctly. The Biserial Correlation indicates that the item discriminated exceptionally well. At the same time, the Item Reliability Index shows that one can depend upon the Biserial Correlation with a high degree of confidence.

The remaining data indicate that very few upper students chose distractors, yet each attracted the lower students. Thus all distractors functioned reasonably well, with alternative 3 being the strongest. This is the most literal of all the distractors and obviously was extremely attractive to the weak students. Apparently lower students, tending to accept the most literal meaning, found it difficult to interpret the selection and grasp the central thought.

The entire item functioned exceptionally well, and might be used as a prototype to teach weaker students how to interpret varying literary selections.

## EVALUATION 4 -- TEST ITEM 93

Read the following quotation and answer question 93.

"Some books are to be tasted, others to be swallowed, and some few to be chewed and digested."

93. "to be tasted" means to be

- A. studied carefully
- B. sampled
- C. disregarded
- D. skimmed

ITEM	N	OMIT	DIF	NR	NF	K	1	2	3	(4)
93	1026	1	0.496		23	4	69	426	10	497
	188	UPPER	5	0			2	76	0	110
	210	UPPER	4	0			5	99	3	103
	214	UPPER	3	8			11	82	1	112
	196	UPPER	2	4			17	82	0	93
	218	UPPER	1	12			34	87	6	79
TEST SCORE MEANS							38.8	48.5	37.0	49.9
Z-SCORE MEANS							-0.85	0.07	-1.01	0.15
BISERIAL CORREL #							0.195			
ITEM REL. INDEX #							0.098			

This item was placed in the comprehension category because the student must translate an abstraction into literal terms.

As approximately 50% of the students selected the correct answer, the item was of acceptable difficulty. The Biserial Correlation, however, is not acceptable, indicating that the item did not discriminate well between upper and lower students.

The Keyed Answer (alternative 4) and distractor 1 show good discrimination between upper and lower students. However, distractor 2 operated poorly, attracting more of the upper students than lower students. Perhaps students selected "sampled" as a better literal translation than "skimmed". As this choice can be defended, possibly the item has no best answer. If a new word, less closely related in meaning to "skimmed" were substituted for "sampled", this item would be strengthened. Distractor 3, almost totally ineffective, must be made more plausible to lower students.\*

\* The suggested revisions (if successful) would produce the following 'chain reaction' effect on the statistics when the new version is tested.

1) If distractor 2 were changed to make it more obviously wrong to upper students, yet still plausible to lower students, more upper students than lower students would be directed away from it. The discriminatory power of distractor 2 would then be increased. Further, most of the upper students formerly choosing 2 would change to the keyed answer, thus making alternative 4 more discriminating.

2) If distractor 3 were changed to make it appear more plausible to only the lower students, then the discriminatory power of this distractor would be increased. Many of these lower students would be drawn from those formerly choosing the keyed answer, thus making alternative 4 even more discriminating.

3) Improving the discriminatory power of these three alternatives would improve the discriminatory power of the entire item and substantially increase the Biserial Correlation and Item Reliability Index.

## EVALUATION 5 -- TEST ITEM 17

17. "All honor to him who shall win the prize,"  
The world has cried for a thousand years;  
But to him who tries and fails and dies  
I give great honor and glory and tears.

Which of the following selections from the text expresses a similar thought?

- A. The Memorial Cup Series  
B. The Service  
C. At the Cedars  
D. Skeleton in Armour

\*

ITEM	N	OMIT	DIF	NR	NF	K	1	(2)	3	4
17	1026	204	0.534		0	2	81	548	44	149
	188	UPPER	5	30			2	141	7	8
	210	UPPER	4	34			9	143	3	21
	214	UPPER	3	61			16	103	9	25
	196	UPPER	2	37			23	90	11	35
	218	UPPER	1	42			31	71	14	60
TEST SCORE MEANS							40.8	51.3	43.9	42.4
Z-SCORE MEANS							-0.66	0.24	-0.36	-0.54
BISERIAL CORREL #							0.385			
ITEM REL. INDEX #							0.192			

To answer the item correctly, the student must first comprehend the given passage, analyze it for theme, and then relate its theme to a textbook selection. This item was placed in the Higher Mental Processes category.

The level of difficulty indicates that slightly more than half the students selected the correct answer. In this respect the item was acceptable. The Biserial Correlation indicates that the item discriminated quite well since those who responded correctly scored higher on the total test. Those who responded incorrectly scored lower on the total test. The Item Reliability Index indicates that one can depend on this Biserial Correlation with a fairly high degree of confidence.

The distractors functioned quite effectively with distractor 4 being the most successful. Students who chose this alternative failed to analyze the passage for theme and probably responded to surface words such as "for a thousand years". The weakest of the distractors is alternative 3. This alternative was probably rejected because At the Cedars contains no mention of a contest and this idea runs throughout the given selection.

Basically the item functioned well, but it could be improved. Alternative 3 could name a selection containing the idea of a contest (e.g., The Diving Fool).

\* To facilitate identification, unacceptable elements are boxed.



## EVALUATION 6 -- TEST ITEM 63

Read the three stanzas below and answer question 63.

- A. To startled skies the pibroch sounds,  
The frenzied chargers champ and neigh;  
With fearful impact battle joins:  
And stark death rides the wind today.
- B. The blaring tocsin calls to battle,  
Claymore thumps on claymore;  
Shot and shell on helmets rattle  
Like pebbles on the sea-shore: ...
- C. The shattering trumpet shrilleth high,  
The nard brands shiver on the steel,  
The splintered spear-shafts crack and fly,  
And horse and rider reel: ...

63. In which stanza are the images most confused?

ITEM	N	OMIT	DIF	NR	NF	K	①	2	3	4
63	1026	8	0.368		0	1	378	354	271	15
	188	UPPER	5	0			63	59	66	0
	210	UPPER	4	1			84	66	58	1
	214	UPPER	3	1			86	67	55	5
	196	UPPER	2	2			72	73	43	6
	218	UPPER	1	4			73	89	49	3
TEST SCORE MEANS							48.5	47.1	49.7	43.2
Z-SCORE MEANS							-0.01	-0.10	0.15	-0.45
BISERIAL CORREL #							0.030			
ITEM REL. INDEX #							0.015			

This item has been classified as testing the Higher Mental Processes. The student is asked to analyze and evaluate three stanzas and decide in which stanza the images are most confused.

The level of difficulty indicates that the students found this item difficult. The Biserial Correlation is not acceptable. The item does not discriminate well between upper and lower students; that is, lower students tended to answer this item correctly as often as the upper students. Since the Biserial Correlation is unacceptable the Item Reliability Index is to be disregarded.

There are several weaknesses in this item. In the stem, the word "confused" was probably misinterpreted to mean something other than disparate or unrelated which the examiners evidently intended it to mean. The attractiveness of the distractors, especially alternative 3, would indicate that many of the upper students accepted the word "confused" as referring to a physical scene rather than to a number of unrelated images. If the word "unrelated" were substituted for "confused", the item should prove more useful.

The vocabulary in the stanzas which served as the alternatives was probably too difficult for Grade IX students. Such terms as "pibroch" and "claymore" are beyond the vocabulary of the average student in this group. Material containing vocabulary of such complexity must be used with caution.

The use of only three alternatives caused difficulty for some students. This may have been avoided had the word "three" in the directions been capitalized, or if the stem had read "In which of the THREE stanzas...."

## EVALUATION 7 -- TEST ITEM 27

(Evaluations 7, 8, and 9 deal with a block of items -- 27, 28, and 29 -- all based on the poem given on this page).

The following poem applies to questions 27 - 30.

A noiseless patient spider,  
 I marked where on a little  
   promontory it stood isolated,  
 Mark'd how to explore the vacant  
   vast surrounding,  
 It launched forth filament, filament,  
   filament, out of itself.  
 Ever unreeling them, ever tirelessly  
   speeding them.  
 And you O my soul where you stand,  
 Surrounded, detached, in measureless  
   oceans of space,  
 Ceaselessly musing, venturing, throw-  
   ing, seeking the spheres to  
   connect them.  
 Till the bridge you will need be form'd  
   till the ductile anchor hold,  
 Till the gossamer thread you fling  
   catch somewhere,  
 O my soul.

27. The above poem is BEST classified as

- A. narrative
- B. dramatic
- C. free verse
- D. ode

ITEM	N	OMIT	DIF	NR	NF	K	1	2	3	4
27	1026	5	0.619		0	3	115	171	635	100
	188	UPPER	5	0			9	?	{31 ? {124	24
	210	UPPER	4	0			23	?	{34 ? {124	29
	214	UPPER	3	0			14	38	147	15
	196	UPPER	2	1			32	?	{33 ? {116	14
	218	UPPER	1	4			37	?	{35 ? {124	18
TEST SCORE MEANS							43.7	47.7	48.8	50.5
Z-SCORE MEANS							-0.36	-0.01	0.07	0.24
BISERIAL CORREL #							0.101			
ITEM REL. INDEX #							0.049			

This item, designed to test familiarity with the forms of poetry, required the student to recall facts about the form of free verse. 62% of the students answered this Knowledge item correctly. Although this level of difficulty is acceptable, the Biserial Correlation indicates that serious flaws exist in the item.

## (Evaluation 7 - cont'd)

The correct alternative was ineffective since virtually as many lower students as upper students answered correctly. The test score mean (average) of those selecting this correct alternative is lower than the test score mean of those choosing distractor 4. Too many upper students were also attracted to distractor 2.

This situation could have been caused by the format\* of the poem. Further the word "O" could have been accepted as a false clue to distractors 2 and 4. It may be, however, that more explicit definitions of the terms used in describing various forms of poetry should become the teacher's concern.

## EVALUATION 8 -- TEST ITEM 28

28. What is compared to the patient spider?

- A. a bridge and anchor
- B. the human soul
- C. gossamer thread
- D. measureless oceans of space

ITEM	N	OMIT	DIF	NR	NF	K	1	(2)	3	4
28	1026	5	0.739		0	2	108	758	35	120
	188	UPPER	5	0			10	173	0	5
	210	UPPER	4	1			10	187	2	10
	214	UPPER	3	0			22	168	5	19
	196	UPPER	2	1			36	120	9	30
	218	UPPER	1	3			30	110	19	56
TEST SCORE MEANS							43.2	50.6	37.8	40.8
Z-SCORE MEANS							-0.45	0.24	-0.93	-0.65
BISERIAL CORREL #							0.508			
ITEM REL. INDEX #							0.223			

To answer the item correctly, the student is required to comprehend the significance of the words of a poem with reference to their context in order to grasp the central thought of the poem. He must perceive a relationship between a spider patiently spinning his web and a figurative description of the soul groping in the world of human experience. The item was placed in the Higher Mental Processes category.

As 74% of the students selected the correct answer, the item was very easy. The Biserial Correlation indicates that the item discriminated exceptionally well between upper and lower students.

Each alternative also discriminated well. However, distractor 3 was not entirely acceptable, attracting less than 5% of the students. To improve the item, this alternative should be made more attractive to more lower students. This revision would make the item somewhat more difficult, and thus decrease the difficulty index.

\* The format of the poem used for this block of items results in a cramped appearance and an unnatural breaking of the poetic lines. This is due to the double columned format of the examination booklet.



## EVALUATION 9 -- TEST ITEM 29

29. The poem illustrates a single sustained

- A. simile
- B. hyperbole
- C. personification
- D. metaphor

ITEM	N	OMIT	DIF	NR	NF	K	1	2	3	(4)
29	1026	15	0.389		0	4	102	138	372	399
	188	UPPER	5	1			21	13	45	108
	210	UPPER	4	4			20	19	75	92
	214	UPPER	3	3			20	29	85	77
	196	UPPER	2	3			19	31	80	63
	218	UPPER	1	4			22	46	87	59
TEST SCORE MEANS							48.0	44.3	46.7	51.0
Z-SCORE MEANS							-0.01	-0.36	-0.10	0.24
BISERIAL CORREL #							0.267			
ITEM REL. INDEX #							0.130			

The item requires a knowledge of terminology and the application of this knowledge to a passage. Analysis of the poem is also required to perceive a relationship between a spider and the human soul as suggested by the poet. The item was placed in the Higher Mental Processes category.

The difficulty level indicates that the item proved to be difficult. The Biserial Correlation indicates that the item did not discriminate well between upper and lower students. Therefore the item is unacceptable.

The weakness of the item becomes apparent when the alternatives are examined. Distractor 1 is of little value as it attracted an equal number of upper and lower students. Distractor 3 is too attractive to upper students. The low Biserial Correlation reflects the low discriminatory powers of these alternatives.

In its present form this item is not effective. To be more acceptable distractors might be revised by substituting all alternatives with descriptive phrases.

## EVALUATION 10 -- TEST ITEM 97

97. The air is still and the lake calm.

Which of the following sentences says this most effectively, the words reinforcing the sense?

- A. I hear lake waters lisp as they lazily strike the shore.
- B. I hear lake water lapping with low sounds by the shore.
- C. I hear the waters of the lake washing on the shore faintly.
- D. I hear the lake's waters splashing on the still shore.

ITEM	N	OMIT	DIF	NR	NF	K	(1)	2	3	4
97	1026	39	0.519		0	1	533	172	207	75
	188	UPPER	5	2			141	16	26	3
	210	UPPER	4	2			132	30	37	9
	214	UPPER	3	12			123	25	45	9
	196	UPPER	2	5			88	44	40	19
	218	UPPER	1	18			49	57	59	35
TEST SCORE MEANS							51.9	44.4	46.0	40.3
Z-SCORE MEANS							0.33	-0.36	-0.20	-0.75
BISERIAL CORREL #							0.445			
ITEM REL. INDEX #							0.222			

It may be maintained that this item tests simply an awareness of aesthetic elements in literature. However, the item does require the student to evaluate in terms of internal evidence since he must decide in which sentence the words most effectively reinforce the desired sense. Thus the item has been placed in the Higher Mental Processes category.

The item is of average difficulty and discriminates well. All the alternatives work very effectively. These data indicate that this is an excellent item.

## COMMENTS

1. In item 63 several students marked a fourth alternative when the item contained only three possible alternatives. Teachers should caution students against assuming that there is a set number of alternatives in multiple-choice test items. While four alternatives are most common with multiple-choice items, three or five are also used.
2. In item 93, it can be argued that there is no one correct answer. This was evident from the analysis when many upper students found an incorrect alternative very attractive. There must always be a definitely best answer to each item and the possibility of dispute should be held to a minimum. The teacher must also be certain that vocabulary is clear and concise and not misleading to the student.
3. It is well to keep in mind that unacceptable statistics need not always be the result of a poorly constructed item. When the analysis indicates that an item has not operated well, the teacher must consider all the following areas when determining the source of the problem:
  - a) technical quality of the item (of the stem and any materials upon which the item is based, as well as of the individual alternatives)
  - b) appropriateness and significance of the concept that he is testing
  - c) instructional procedures that he has used in teaching the concept
4. Many teachers have had little training in the field of statistics. For this reason, the present booklet purposely avoids technical explanations that may be confusing. One evident example of this is found in the explanation of the Item Reliability Index. Completely disregarding the source from which this index is derived, only the most practical consequence of its use in interpreting an item analysis is stated in this booklet (see CHART 3).

5. In testing Literature IX, the Knowledge level (memorization) is often stressed. Present practices, however, indicate a shift of emphasis from simple recall and a preoccupation with specific works to the student's ability to understand and use literary concepts. Releases from the Department of Education indicate that the 1968 Grade IX Departmental Examination will not include items from the text and that 80% of the test will focus on items testing Higher Mental Processes beyond the Knowledge level. This, however, should not be misconstrued to mean that the Knowledge level is to be neglected. Rather, it is to be considered as a basis for the development of abilities in understanding and using literary concepts at a higher mental level.
6. Computer services may not be readily available to many teachers. Where this is the case these teachers may wish to obtain the pamphlet, Short-cut Statistics for Teacher-made Tests, published by the Educational Testing Service, Berkeley, California. The teacher is shown how he may, within his own classroom situation, roughly approximate some of the information obtainable from the computer analysis dealt with in the present booklet.
7. Periodically the computer program for item analysis is revised to include additional information or slightly change the format. For this reason future computer 'print-outs' may vary slightly from the ones shown in this booklet. Recent changes have included the addition of the Item Reliability Index (incorporated in the analyses appearing throughout this booklet), and the total number of students in each of the five groups as indicated in the column under 'N' of Charts 2, 5, and 6. A current change (not shown in this booklet) is the use of the word 'GROUP' to replace the word 'UPPER' as it appears in all the foregoing item analyses.



## REFERENCES

For a further study of testing principles teachers may wish to refer to materials listed below. Educational Testing Service has also published an excellent pamphlet listing further references, each accompanied by a brief descriptive abstract (see Locating Information on Educational Measurement: Sources and References).

Anastasi, Anne, Editor, Testing Problems in Perspective, Washington, D.C.: American Council on Education, 1966.

Ayers, J. D., Test Item Construction: A Manual For Teachers, Edmonton, Barnett House, 1967.

Bloom, B. S. (et al), Taxonomy of Educational Objectives, Handbook I: Cognitive Domain, New York: David McKay Company, Inc., 1956.

College Entrance Examination Board, End-Of-Year Examinations in English for College-bound Students grades 9-12, Princeton, New Jersey, 1963.

Department of Education, Summary Description of Grade Nine Literature Objectives, Test Items and Blueprint, Edmonton, 1967.

Department of Education, Summary Description of Grade Nine Science Objectives and Items, Edmonton, 1965.

Department of Education, Summary Description of Grade Nine Social Studies Objectives and Items, Edmonton, 1966.

Department of Education, Supplement to Summary Description of Grade Nine Science Objectives and Test Items, Edmonton, 1966.

Department of Education, Supplement to Summary Description of Grade Nine Social Studies Objectives, Test Items, and Blueprint, Edmonton, 1967.

Department of Education, Taxonomy of Language IX Objectives with Illustrative Test Items: A Summary Description, Edmonton, 1968.

Department of Education, Taxonomy of Mathematics IX Objectives with Illustrative Test Items: A Summary Description, Edmonton, 1968.

Downie, N. M., Fundamentals of Measurement (2nd edition), Toronto: Oxford University Press, 1967.

Ebel, R. L., Measuring Educational Achievement, Englewood Cliffs, N.J.: Prentice Hall, 1965.

Educational Testing Service, Locating Information on Educational Measurement: Sources and References, Berkeley, California, 1965.

Educational Testing Service, Making the Classroom Test - a Guide for Teachers (2nd edition), Princeton, New Jersey, 1961.

Educational Testing Service, Making Your Own Tests: A Work Kit To Accompany Three Film Strips, Princeton, New Jersey, n.d.

Educational Testing Service, Multiple-Choice Questions: A Close Look, Princeton, New Jersey, 1963.

Educational Testing Service, Short-cut Statistics for Teacher-made Tests, Berkeley, California, 1964.

Furst, E. J., Constructing Evaluation Instruments, Toronto: Longmans, Green & Co., 1958.

Gerberick, J. R., Specimen Objective Test Items, Toronto: Longmans, Green & Co., 1956.

Hedges, William D., Testing and Evaluation for the Sciences, Belmont: Wadsworth Publishing Co., Inc., 1966.

Krathwohl, D. R., B. S. Bloom, and B. B. Masia, Taxonomy of Educational Objectives The Classification of Educational Goals, Handbook II: Affective Domain, New York: David McKay Company, Inc., 1964.

Sanders, N. M., Classroom Questions: What Kind?, New York: Harper & Row, 1966.

Thorndike, Robert L., and Elizabeth Hagen, Measurement and Evaluation in Psychology and Education (2nd edition), New York: John Wiley & Sons, Inc., 1961.

Womer, Frank B., One-Year Schoolwide Project Grades K - 6, Chicago: Science Research Associates, Inc., 1966.





[illegible]

PR 87 A33 1968

ALBERTA DEPT OF EDUCATION  
JUNIOR HIGH SCHOOL EXAMINATIONS  
BOARD

39844382 CURR HIST



\*000017904277\*

PR 87 A33 1968

Alberta. Dept. of Education.  
Junior High School Examinations  
Board.

39844382 CURR HIST

CURRICULUM GUIDE

**For Reference**

NOT TO BE TAKEN FROM THIS ROOM  
EXCEPT WITH LIBRARIAN'S PERMISSION

